



Designing a Digital Reading Tool for Middle School Students and Teachers: A Grounded, User-Centered Design Approach

John P. Hutson^(✉), Kinta D. Montilus, Kristen S. Herrick, and Michelle R. Lamond

Educational Testing Service, Princeton, NJ 08541, USA
jhutson@ets.org

Abstract. Across English Language Arts classrooms in the United States many students experience low reading comprehension and struggle to stay engaged with their learning while working independently [1]. Our cross-functional team partnered with middle school teachers and students to develop an educational tool that provides reading skills support. Through three design thinking stages, we present our method, experiences, and initial outcomes developing a reading tool for middle school English Language Arts classrooms. In Stage One, we empathized with users to understand their needs and pain points and identified the problem statement: students often do not work well independently and struggle to stay engaged with reading. In Stage Two, we explored design solutions to address the problem statement and engaged in iterative prototyping. At the end of this stage, we developed an alpha version of our reading tool that included both annotation and reading comprehension check features. In Stage Three, we conducted initial user testing and evaluation of the tool in middle school classrooms (Teacher N = 3; Student N = 171). Teachers found they were able to implement the tool flexibly to fit their planned lessons and instruct students at different reading levels. Students worked independently with the tool, reported that it was usable, and perceived it would improve their reading comprehension. Overall, this study shows how a tool with annotation and metacognitive-check features can support students in the classroom, and offers next steps for the development of effective reading tools.

Keywords: User studies · Education technology · Reading literacy

1 Introduction

Low reading literacy among middle school students is a problem across the United States. In 2019, 66% of 8th grade students read below a proficient level [2], and 27% of those were below the basic level. Given the breadth of the reading skills needed to be a proficient reader and the abundance of work in this area, there are many competing theories, practices, and technologies. We engaged in a grounded, user-centered design approach, and present work leading up to and including an initial implementation of our tool in classrooms.

1.1 Method Overview

Our interdisciplinary team implemented two key approaches to develop a solution to support reading comprehension skills of struggling middle school readers: (1) user-centered design [3] and (2) grounding in science and practice [4]. While both approaches have been used widely, their combination to address reading literacy in the middle school classroom created unique opportunities, challenges, and methodological choices that our team learned from, and we expect others could incorporate into their work.

In Stage One of the design thinking process we engaged in a discovery process by conducting interviews with teachers and students to identify pain points and needs. In Stage Two we explored potential design solutions rooted in learning science and instructional design principles, and conducted frequent user tests to develop an alpha version of the reading tool. Finally, in Stage Three, we conducted an alpha study of the tool in classrooms with teachers and students to evaluate perceptions of usability and preliminary evidence of effectiveness. All studies were approved by IRB and participants completed informed consent (see Stage Three for full study ethics).

2 Stage One: Empathize with Users and Define Problem

The work in stage one aligned with the first step in the HCI Design Process: research and requirements gathering. We conducted surveys ($n = 31$), focus groups ($n = 27$), and journey map interviews ($n = 4$) with middle school English Language Arts (ELA) teachers. We also conducted surveys with students ($n = 47$). The two main outcomes of Stage One were developing an empathy map to understand teacher and student needs and pain points, and using those to develop a problem statement to guide the development of the learning solution.

Through empathy mapping, we established that middle school teachers would benefit from support meeting the needs of learners at different levels, helping learners to get started with their work, helping learners to feel empowered and encouraged to complete work independently, helping learners feel engaged/excited to read, and providing adequate and appropriate resources.

After summarizing the major pain points of middle school ELA teachers and students, we curated six problem statements and tested them with teachers ($n = 3$). Teachers rated how much they identify with the problem statements and edited the problem statements to align it with their actual needs. After this feedback, the team selected the problem statement: *As a general education 7th-grade ELA teacher, I am discouraged when my low reading comprehension students do not work well independently because they struggle to stay engaged with their learning.*

3 Stage Two: Explore Design Solutions and Engage Users in Iterative Prototyping

Equipped with a better understanding of our users' needs and experiences, in Stage Two, we explored and collaboratively ideated on potential solutions. The problem statement

served as our guide throughout this stage. The work in stage two aligned with the second step in the HCI Design Process: Design and Prototyping.

First, we identified and considered applicable learning science research concerning reading comprehension, motivation to read, metacognition, and self-regulated learning. Given that teaching and learning related to middle school literacy is a well-researched area, we wanted to ensure we incorporated insights from applicable learning science research into our brainstorming, design thinking, and ideation work. We also reviewed and incorporated applicable learning design principles (LDPs) and instructional best practices while exploring design solutions.

To ground our work, we adopted the comprehensive Cognitive Based Assessment of, for, and as Learning (CBAL) reading framework [5] as our primary learning framework underpinning design solution exploration. Drawing from the framework, we divided the reading process into the following dimensions: 1) preparing to read, 2) understanding the text, 3) digging deeper or going beyond the text, 4) re-representing text information, and 5) applying and reflecting.



Fig. 1. Screenshots of the ELaborate welcome screen with instructions, the text view with annotation modal window, and the final theme response box.

Having established a strong learning science and instructional design foundation on which to build, we ideated and created rapid prototypes. In Stage Two we regularly conducted studies with students and teachers on our rapid prototypes to evaluate the extent to which design decisions met their needs. The two key features we decided to

focus on in the initial prototype were 1) annotation and 2) reading comprehension self-check questions. Annotating text has positive impacts on student reading comprehension through reading skill, metacognition, and socio-emotional learning [6]. Reading comprehension checks presented with the text have positive impacts on reading comprehension, metacognition, and socio-emotional mechanisms [5].

Through exploration of design solutions and iterative, rapid prototyping, we integrated insights from learning sciences research and instructional design principles with the expertise of our users to create a functional prototype – ELAaborate. In ELAaborate (Fig. 1), students: 1) view a demo video to help them understand how to use the tool as well as the assignment instructions [7], 2) answer a guiding question [8], 3) read the text while annotating via hashtags, highlights, and comments and review their annotations as needed [9], 4) answer reading comprehension check questions [10], and 5) write a brief summary about the story’s theme [11]. ELAaborate had a single text, Thank you M’am by Langston Hughes.

4 Stage Three: Initial Testing and Evaluation

In stage three, we conducted an alpha study of our prototype (ELAaborate) to determine the extent to which it was able to address needs identified in our problem statement. The work in stage three aligns with the third step in the HCI Design Process: evaluating designs.

4.1 Research Questions

Study goal 1 asked how do teachers choose to use the tool? There were two research questions associated with this goal: 1) do teachers implement the tool in similar ways, and 2) why did teachers implement the tool in the way they did? Study goal 2 asked how students and teachers perceive the tool? There were two research questions associated with this goal: 1) do users perceive the tool as usable, and 2) do users perceive the tool as improving reading comprehension? For each research question, the prediction was that there would be supporting evidence to support these questions, because ELAaborate was designed specifically to address these constructs. However, a main purpose of Stage Three is to identify areas for improvement and optimization, so below we also use exploratory analyses to probe when there is negative evidence for a research question.

4.2 Method

The study was conducted in middle school (7th & 8th grade) ELA classrooms. ELAaborate was implemented as a regular classroom activity. There were three teachers in the study, with a total of 9 classes. A total of 171 students participated in at least a portion of the classroom activities. Data collection was remote, and we only had contact with teachers. Teachers chose how to implement the tool into their classroom instructions and additional activities for their class around the reading. Teachers were paid \$100 per hour for participation outside of classroom (e.g., training and interviews). Schools were compensated \$20 per student who completed the study. The full study design was

approved by IRB, and approval was received for each school using their district protocols. All participation was voluntary, and participants consented to participate. Students used anonymous participant IDs assigned by teachers throughout the study, so all student data was deidentified throughout the study. Teacher data was deidentified upon completion of the study.

Students completed multiple surveys in addition to their engagement with the ELAbrate prototype. Before using ELAbrate, students were asked five questions to gauge their motivation, reading habits, and comfort with technology. While reading, students were given the goal of identifying the theme of the story. Use of the annotation feature was optional, but all reading check questions had to be answered correctly before moving on. After using the tool, students completed usability items (System Usability Scale [SUS; 12], and perception of impact items. The perception of impact questions gathered students' agreement (1 = Strongly disagree, 5 = Strongly agree) that ELAbrate supported six outcomes the tool was designed to address: Overall text comprehension, vocabulary, ability to annotate, motivation, engagement, and independent reading. Students also rated agreement (1 = Strongly disagree, 4 = Strongly agree) with two reading assignment specific items: needed support from teacher to understand 1) plot and 2) theme. Teachers participated in follow-up interviews to discuss the implementation of the tool.

For teacher interviews, we conducted a thematic analysis. The implementation themes were flexibility of implementation and logistics of implementation. We present descriptive statistics for the survey items to illustrate the trends in the data. Additionally, we use correlations and Bayes factors as exploratory analyses to probe the data. Finally, where applicable, we present the interview and survey data together to put the survey results in context.

4.3 Results

Study Goal 1: Teacher Implementation. (1) Do teachers implement the tool in similar ways, and (2) why did teachers implement the tool in the way they did? All teachers in this study implemented the tool differently. There were some similarities, such as completing most of the reading independently. Notably, all teachers provided different levels of support and structure during the reading component. As such, the key outcome here is that teachers found the tool flexible to different implementations, and the tool did not create limitations on their implementation plans.

Teachers who gave their students the most freedom to complete the reading independently stated that they wanted to understand how students would naturally use the tool. However, one of these teachers changed their implementation based on the reading level of their students. For example, they varied how much annotation demonstration they did. Another teacher chose to scaffold their students through the first use of the tool in order to increase the likelihood they would understand how to use all the tool features. They also indicated that with future uses they would increase the amount of independent work in ELAbrate.

Study goal 2: Perception of Usability and Impact of Reading Comprehension. (1) Do users perceive the tool as usable? Overall, students gave the tool high usability ratings. This aligned with teachers' perceptions that the tool was easy for students to

use. The System Usability Scale (SUS) score for ELaborate was in the “Good” range ($M = 74.14$; $SD = 15.23$), which is especially high for the first classroom release of ELaborate. SUS scores did not vary between teachers ($BF = .10$). ELaborate usability can still increase, and one feature identified for improvement was that the process to save annotations was cumbersome for certain students (“I believe it would be easier if the quotations automatically save as you type them.”).

(2) Do users perceive the tool as improving reading comprehension? Overall, on the six reading comprehension outcomes completed after using the tool, students perceived that using ELaborate supported their reading comprehension (Fig. 2A), with each measure above the neutral point on the scale (p 's $< .001$). Importantly, these measures were the same across teachers (BF range .47–.11) with one exception. For working independently, one teacher's students perceived the tool as not having a positive or negative impact on working independently ($BF = 1.53$; working independent $M = 3.2$, $SE = .21$). There was not clear evidence to explain this result in the implementation data, but future work should explore this key outcome from the problem statement.

We ran correlations on the six reading comprehension outcomes, and all were positive (p 's $< .05$; min $r = .21$; max $r = .75$). However, annotation is one of the key features of ELaborate, and critically most of the lower correlations include the item that ELaborate would support students in “annotating more.” For example, the lowest correlation ($r = .21$) is between annotating more and “understood text more,” which using Fisher's Z transformation is significantly weaker than the relationship between both vocabulary and motivation with understanding (p 's $< .05$). In other words, students perceived that reading the story in ELaborate supported their comprehension of the story, but that annotation may have had less of a relationship with comprehension improvements than other features in the tool. This appears to be driven by students who indicated the tool allowed for a good understanding of the text (Fig. 2B), such that students who gave high understanding ratings were more likely to give lower annotation ratings. Future research with the tool should explore when annotations support comprehension.

Most students indicated they did not need help to understand the plot ($M = 1.59$) or theme ($M = 1.81$) of the story (p 's $< .001$). However, there was one teacher for whom students perceived they needed more help compared to the other two classes for both theme and plot (BF 's > 1084). Importantly, their means were above the midpoint of the

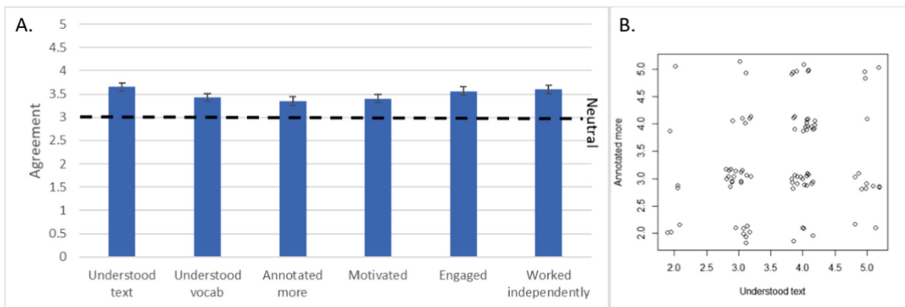


Fig. 2. A. Bar graph of means for the student perceived outcomes items. Error bars are standard error. B. Scatter plot of the outcomes understood text (x-axis) and annotated more (y-axis).

scale (Plot $M = 2.29$; Theme $M = 2.79$) and were a full point higher than the next score on the four-point scale used. One potential reason for this is that this teacher discussed each page of the story with their students before moving on to the next page, which suggests implementation can impact perceptions of the impact of the tool.

5 Discussion

How can we help teachers in middle school ELA classrooms support students who are struggling to read, work independently, and engage in their learning? We developed a prototype reading tool (ELaborate) for use in classrooms through a three-stage research and design process of 1) defining a problem statement, 2) designing a prototype, and 3) testing the solution. In this report, we presented the method and key results/outcomes from an interdisciplinary approach of frequent user testing and grounding solution features in science and practice.

5.1 Implications

This study has implications both specific to ELaborate and learning tools, and to reading broadly. For both teachers and students, it was important that the tool gave them flexibility in implementation to fit their needs. However, the way the tool is implemented is likely to impact its effectiveness, which means it is important for the development team to understand this relationship and communicate it to teachers to support their implementation and/or create features in the system to support students while working independently. Also, students perceived that ELaborate's features of annotation and comprehension check questions supported their comprehension overall. However, the perceived impact of annotation was weaker than some other components. As such, future work will need to identify the unique impact of annotation features above and beyond other potential areas for development.

5.2 Conclusion

Our team developed a unique solution to a highly complex problem (low reading literacy) in a context with multiple users (students & teachers). Through the discovery phase we identified annotation and reading comprehension checks as key features to support student literacy. Initial testing of these features in ELaborate was positive, and identified future development and research needed to create an effective reading solution. Our approach to this work had aspects that were unique to the classroom context that we hope other researchers focusing on classroom challenges can learn from.

References

1. Sabatini, J., O'Reilly, T., Doorey, N.A.: *Retooling Literacy Education for the 21st Century: Key Findings of the Reading for Understanding Initiative and Their Implications*. Educational Testing Service, 2018. Center for Research on Human Capital and Education

2. Institute of Education Sciences, National Center for Education Statistics, U.S. Department of Education (ed.), *National Assessment of Educational Progress* (2019)
3. McWilliams, K., et al.: Unpacking the black box of efficacy: a framework for evaluating the effectiveness and researching the impact of digital learning tools [White Paper]. In: *MacMillan Learning* (2017)
4. Graesser, A.C., et al.: Educational technologies that support reading comprehension for adults who have low literacy skills. In: *The Wiley Handbook of Adult Literacy*, pp. 471–493 (2019)
5. O'Reilly, T., Sheehan, K.M.: Cognitively based assessment of, for, and as learning: a framework for assessing reading competency. *ETS Res. Rep. Ser.* **2009**(2), i–43 (2009)
6. Johnson, T.E., Archibald, T.N., Tenenbaum, G.: Individual and team annotation effects on students' reading comprehension, critical thinking, and meta-cognitive skills. *Comput. Hum. Behav.* **26**(6), 1496–1507 (2010)
7. Collins, A., Kapur, M.: Cognitive apprenticeship. In: Sawyer, R.K. (ed.) *The Cambridge Handbook of the Learning Sciences*, pp. 109–127. Cambridge University Press, Cambridge (2014)
8. Rosenshine, B., Meister, C., Chapman, S.: Teaching students to generate questions: a review of the intervention studies. *Rev. Educ. Res.* **66**(2), 181–221 (1996)
9. AbuSeileek, A.F.: Hypermedia annotation presentation: the effect of location and type on the EFL learners' achievement in reading comprehension and vocabulary acquisition. *Comput. Educ.* **57**(1), 1281–1291 (2011)
10. Agarwal, P.K., Nunes, L.D., Blunt, J.R.: Retrieval practice consistently benefits student learning: a systematic review of applied research in schools and classrooms. *Educ. Psychol. Rev.* **33**(4), 1409–1453 (2021)
11. Graham, S., Hebert, M.: Writing to read: evidence for how writing can improve reading. In: A report from Carnegie Corporation of New York. 2010, Alliance for Excellent Education: Washington, DC
12. Sauro, J., Lewis, J.R.: When designing usability questionnaires, does it hurt to be positive? In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011)